ImageTexts

Studying Images and Texts in Conjunction

Melvin Wevers, DH Lab, KNAW HUC, The Netherlands
Thomas Smits, Radboud University, The Netherlands
Leonardo Impett, EFPL, Switzerland

Research Question

The recent upsurge of large-scale analysis of visual material (Computer Vision) shifts the focus in Digital Humanities research away from texts. However, this has also led researchers to approach text and images as disjointed entities. **We analyze similarity and change in both textual and visual elements of car advertisements extracted from digitized newspapers.** By juxtaposing change over time in text and visual material, we aspire to show that the meaning of imagetexts can be studied by looking at the relation between the two forms of representation.

Dataset

Our dataset consists of 9,863 the advertisements for cars extracted from the Dutch newspaper *De Volkskrant* between 1945 and 1995. These advertisements have a visual and a textual component (see Fig. 1).

Fig. 1: Automobile advertisements from *Volkskrant*

Methods

Text: Kleinberg's (2002) burst algorithm to detect 'bursty' words in streams of text data extracted from advertisements. The algorithm uses a probabilistic automaton that identifies state transitions that correspond to points in which the frequency of words changes significantly. (Figs. 2, 3, & 4) We used the Prophet algorithm (Taylor, 2017), an additive regression model, to detect change points in the relative monthly burstiness. (Fig . 2)

Images: Generative Adversarial Networks, a system of two convolutional neural networks, gradually learn from each other, until the synthesis network produces images indistinguishable from those of the original dataset. The latent subspace of generative adversarial networks bears remarkable semantic information, allowing the possibility for semantic vector arithmetic similar to that found in the Word2Vec word embedding model (Goodfellow et al., 2014). The verisimilitude of the generated images is an indication of the meaningfulness of the learned subspace.

Results: texts

Burstiness in ads increased and ten change points could be identified (see fig. 2). Periods of burstiness correspond to particular themes, such as fuel efficiency, environment, safety, and gadgets (see fig. 3)

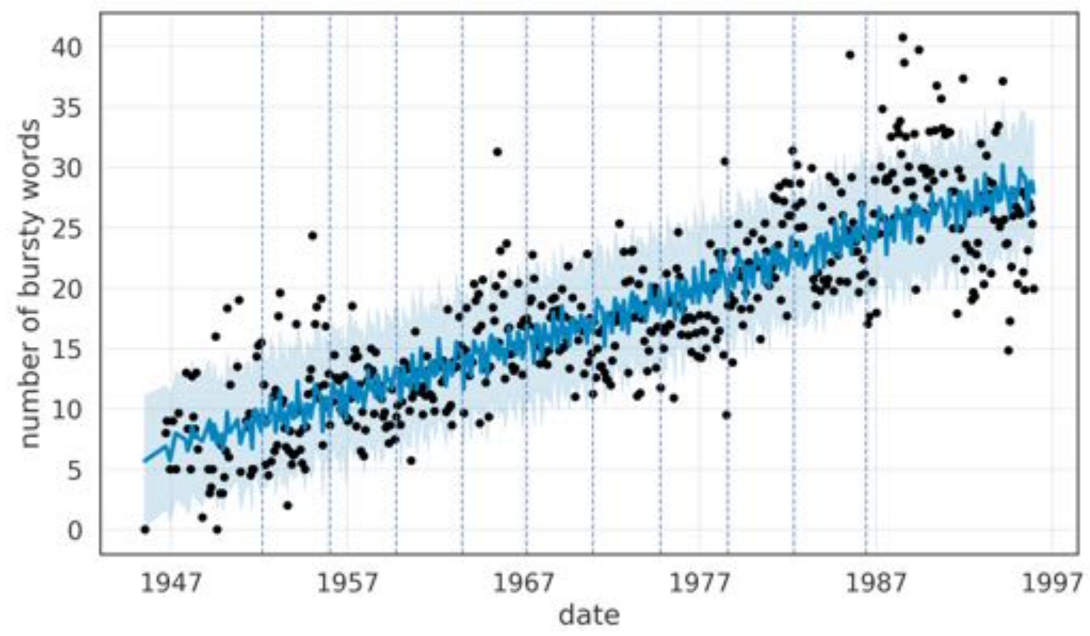

Fig. 2: Average Monthly Burstiness and change points in burstiness

50s-60s	70s	80s	90s
Heating	Fuel efficient	Turbo	Airbag
Cooling	Comfort	Injection engine	Speaker system
Disc brakes	Credit	Catalytic convertor	Seatbelt tensioner
Radiator fan	Financing	Lambda sensor	Waste disposal fee

Fig. 3: Selection of Bursty words in the five decades of the dataset

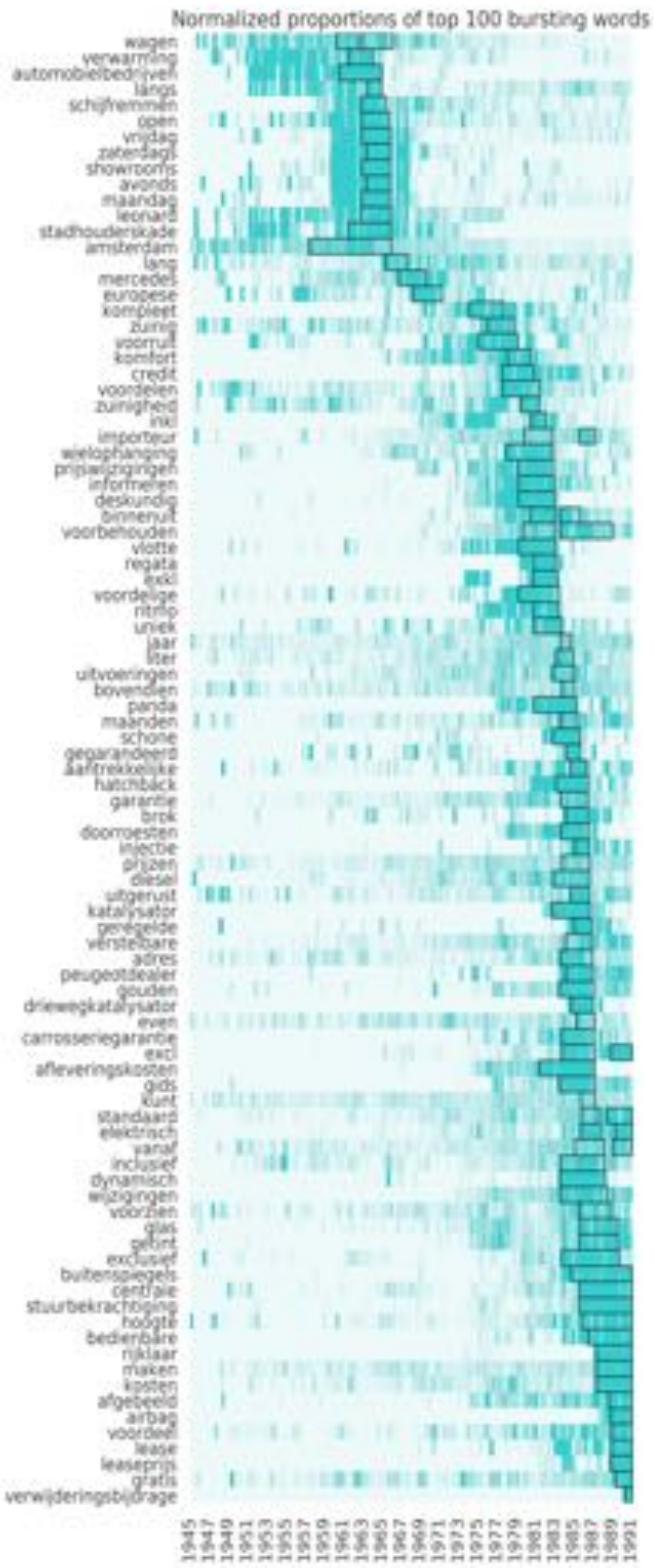

Fig. 4 Heatmap of top 100 bursty words

Results: images

GANS were unable to faithfully reproduce the range of styles and composition present across five decades of automotive adverts. We were able to isolate a large number of car images, warped to a standard size. With this more consistent set of images, a GAN was able to learn the variances in car models, styling, color, position and photographic composition seen in the adverts themselves (see figures below).

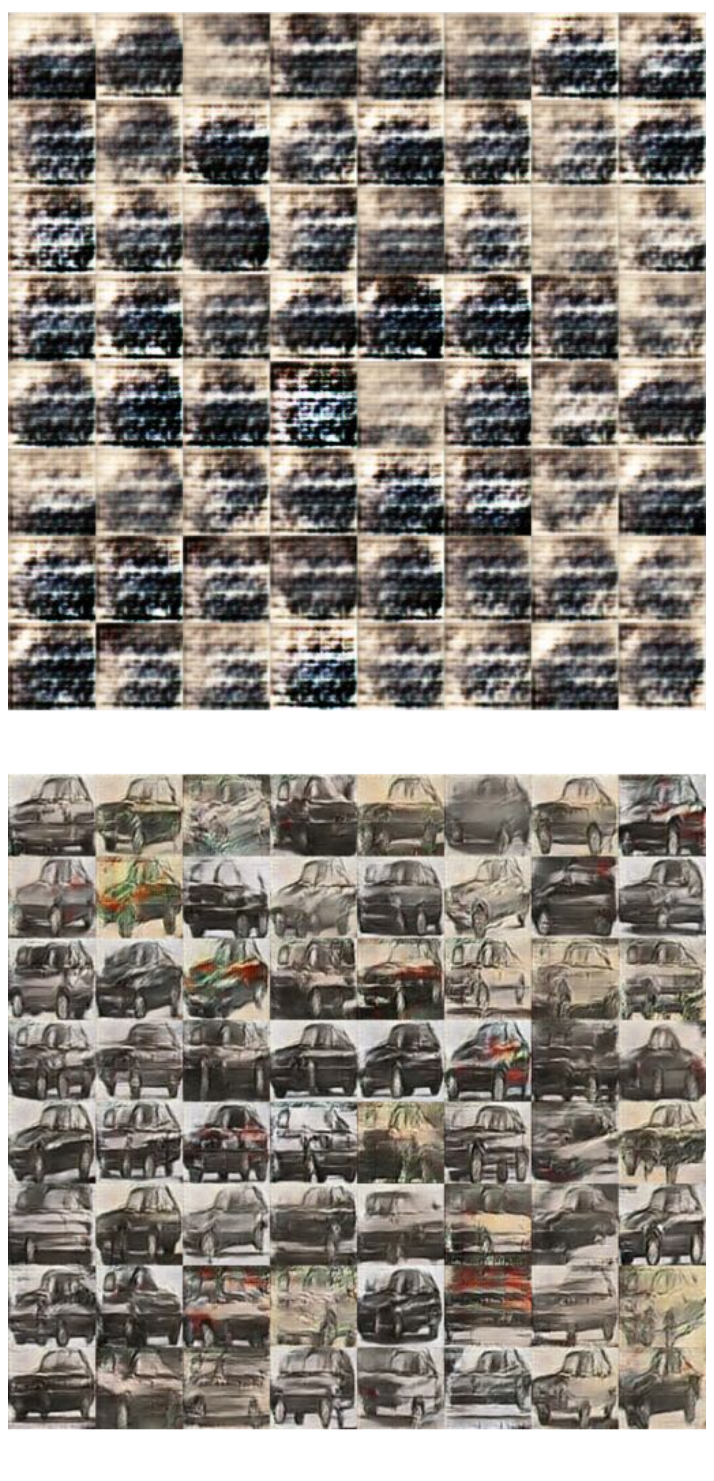

Fig. 4: GAN after 1 & 57 epochs

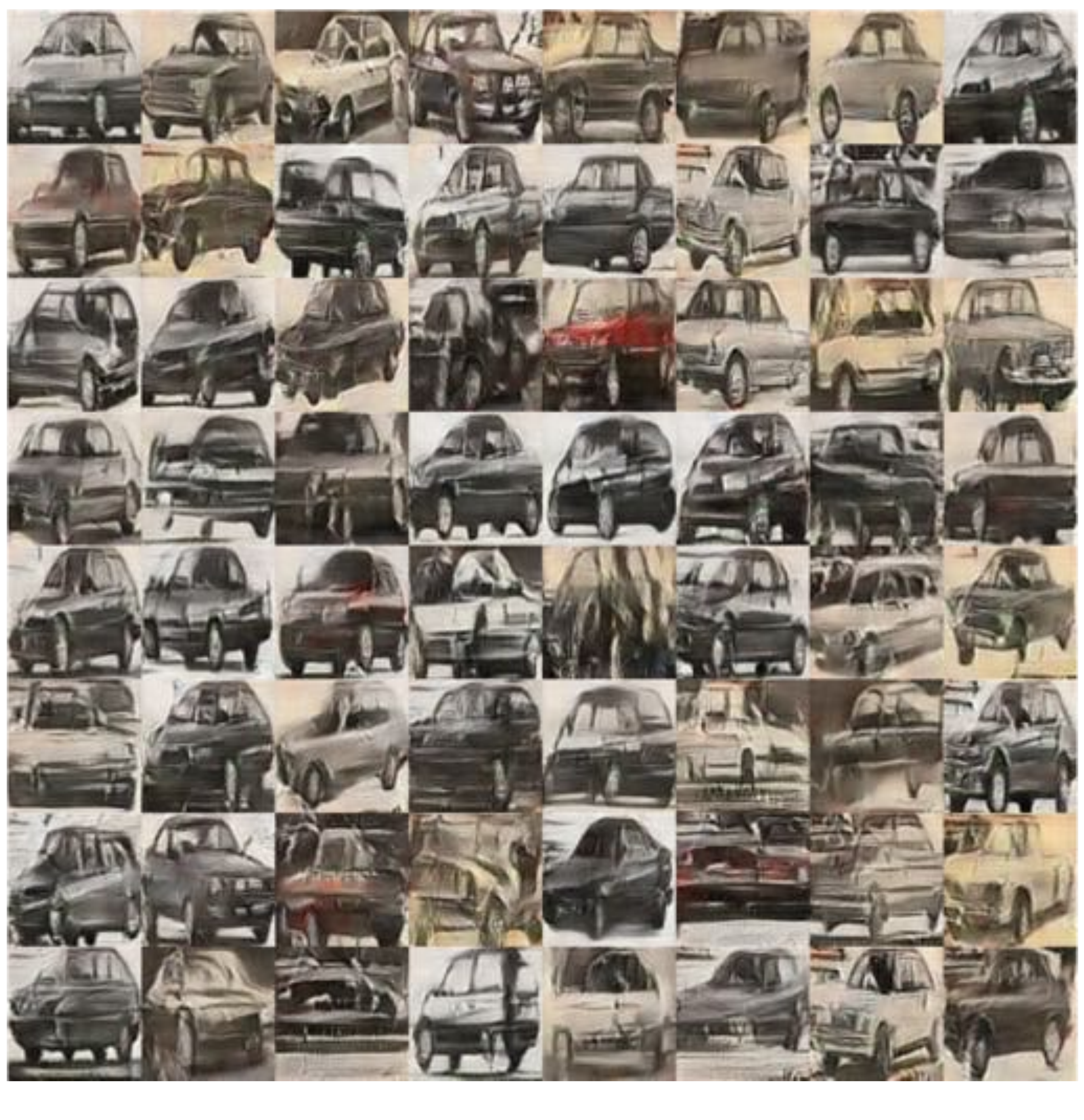

Fig. 5: GAN after 249 epochs

References

- Goodfellow, Ian J., Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. "Generative Adversarial Networks." *ArXiv:1406.2661 [Cs, Stat]*, June 10, 2014.
- Kleinberg, Jon. "Bursty and Hierarchical Structure in Streams." In *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 91–101. KDD '02. New York, NY, USA: ACM, 2002.
- Taylor, Sean J., and Benjamin Letham. "Forecasting at Scale." PeerJ Inc., September 27, 2017.